

Article

Not peer-reviewed version

PREDICTIVE MODELING OF CO₂ METHANATION REACTION FOR CARBON CAPTURE

[Osman Okur](#)^{*}

Posted Date: 8 September 2023

doi: 10.20944/preprints202309.0539.v1

Keywords: methane, carbon dioxide, modelling, artificial intelligence



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Predictive Modeling of CO₂ Methanation Reaction for Carbon Capture

Osman Okur

TUBITAK Marmara Research Center, Gebze, Kocaeli, 41469, Turkey

Abstract: Among various processes, catalytic CO₂ methanation has emerged as a promising method for carbon capture and utilization. Therefore, a CO₂ methanation reaction studied with experimentally. The extent of coke formation on the catalyst after the reaction was quantitatively studied. Then, four different machine learning model were employed in order to model CO₂ methanation process and to predict CO₂, CO, CH₄, H₂ and N₂ concentrations at outlet. The results of the machine learning models were consistent with the experimental data. Performances of the proposed regression methods are evaluated with 10-fold cross validation. Random forest and decision tree regression performed the best among other methods by achieving R >0.98 for the output concentrations and were able to outperform other modeling approaches. It was shown that the best result among the literature studies is reached in this study. The results have demonstrated the possibility of simulation of the methanation process via machine learning techniques. These techniques can be used control and optimization of carbon capture.

Keywords: methane; carbon dioxide; modelling; artificial intelligence

1. Introduction

Reducing carbon dioxide (CO₂) emissions is critical to addressing the global challenge of climate change. Among various processes, catalytic CO₂ methanation has emerged as a promising method for carbon capture and utilization [1]. To address global warming, current policies encourage the use of carbon capture and storage (CCS) technologies. Properly controlled and optimized CO₂ methanation reactions can significantly contribute to carbon sequestration and reduce environmental impact. The carbon dioxide (CO₂) methanation process is an important by-product of carbon capture and utilization technologies and holds promise for renewable natural gas (SNG) production. This process converts CO₂ and hydrogen (H₂) to methane (CH₄), which is a valuable energy carrier and chemical precursor [2]. Power and gas technologies enable cooperation in energy sectors such as electricity, gas, heating and transportation. It also offers economic and environmental benefits to the energy sector due to its efficiency. In addition to capturing greenhouse gases and injecting them into the economy, they also use other forms of renewable energy. Methane (CH₄), a valuable fuel and food substitute, reacts to form CO₂ and hydrogen with ready transportation and distribution systems. Adding renewable energy to methane production increases the environmental impact of the process. Carbon dioxide methanation is also a way to store excess renewable energy and transport it through the existing natural gas infrastructure. However, optimization of the methanation process for large-scale industrial applications is important. As research in this area continues, further advances in catalyst development and process optimization are expected, paving the way for more sustainable carbon management strategies and contributing to a more carbon neutral energy system [3]. Machine learning techniques have been increasingly employed to analyze and predict CO₂ conversion in this process. Yılmaz et al. [1] utilized a comprehensive dataset to develop a random forest model, successfully predicting CO₂ conversion based on various catalyst properties and reaction conditions. Similarly, Dashti et al. [4] investigated CO₂ methanation, incorporating metallic promoters to enhance catalyst performance. Their genetic programming model effectively estimated CO₂ conversion and CH₄ selectivity.

Further advancements in CO₂ hydrogenation catalysts have been explored through data science and machine learning. Fedorov and Linke [5] collected and analyzed CO₂ hydrogenation catalyst data using the Anderson-Schulz-Flory distribution, revealing key factors affecting selectivity. Ngo and Lim [6] proposed a physics informed neural network model for simulation of fixed bed reactor, achieving accurate CO₂ methanation predictions. Suvarna et al. [7] presented a machine learning framework for predicting methanol yield, highlighting the significance of descriptors like metal content and space velocity.

Theory-guided machine learning has been applied to noble metal catalyst performance in the water-gas shift reaction [8]. This approach integrates physical and chemical properties to achieve accurate predictions while adhering to thermodynamic principles. Pérez et al. [9] introduced a novel microstructured reactor design, offering improved performance in exothermic reactions like the Sabatier process. Shen et al. [10] focused on methanation optimization, developing a supervised learning based on an extreme learning machine algorithm, demonstrating accurate CO₂ conversion rate predictions.

Furthermore, machine learning has been employed for adsorbate binding energy predictions on metal surfaces [11]. Praveen and Comas-Vives [11] designed a unified algorithm to forecast various adsorbates binding energies on different facets of transition metals, with potential for broader applications.

In this study, we delve into the predictive modeling of CO₂ methanation reactions for carbon capture applications. By utilizing a dataset of experimental results of CO₂ methanation study, we aim to develop a robust predictive model that can effectively predict CO₂, CO, CH₄, H₂ ve N₂ concentrations. Our approach will take into account a number of influencing factors such as gas velocities, reaction temperatures, pressures, and H₂/CO₂ ratios. The developed predictive model holds the potential to guide the design and optimization of catalytic systems for CO₂ methanation, ultimately contributing to more efficient and sustainable carbon capture strategies.

2. Data Collection and Experimental Procedure

The methanation test set-up for the production of methane, a valuable product resulting from the capture and use of CO₂, currently the most promising and popular technology, is shown in Figure 1. The input gases required for methane production from CO₂ are CO₂ and H₂.

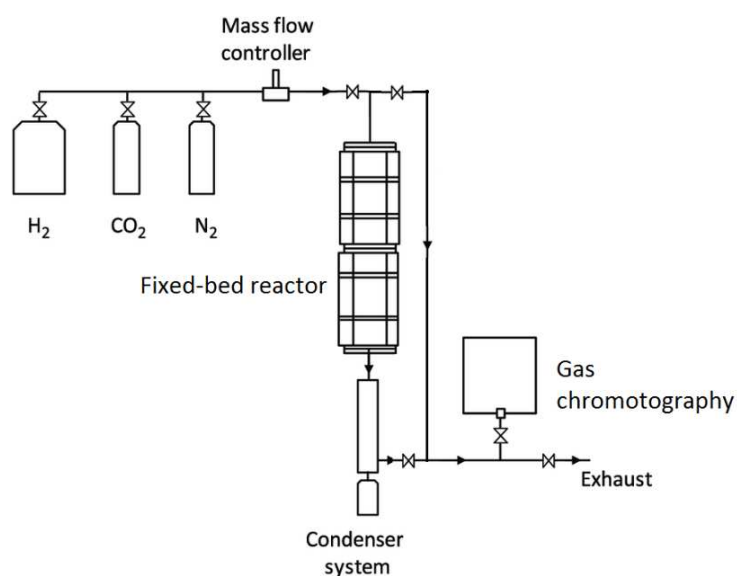
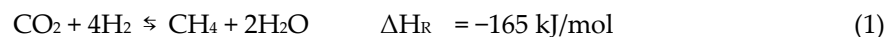


Figure 1. Schematic diagram of the CO₂ methanation process.

The methanation of CO₂ is an exothermic catalytic reaction and is typically operated at temperatures between 250°C and 400°C [12] depending on the catalyst used:



The methanation reaction was performed at experimental set-up in a fixed bed stainless steel 216 reactor with an inner diameter of 10 mm shown in Figure 2. This reactor was fixed upright inside Nabertherm NB300 furnace and a thermocouple was embedded into the center of catalyst bed to measure the temperature during the reaction.

Catalyst testing for activity were performed by using 0.5 grams of catalyst particles with 50–150 microns. Then 0.5 g of quartz particles of the same size range were added in equivalent volume to avoid localized heating. The desired gas compositions from ALICAT MFCs were adjusted and fed to the system according to the reaction conditions. Before the reaction, calcined catalysts were treated for reduction with a gas mixture composed of 10v% H_2 -90v % N_2 at 350 °C for 4 h under atmospheric pressure [13].



Figure 2. CO₂ Methanation experimental set-up.

The product gas was quantitatively analyzed every 30 min by an on-line (Agilent, 7820A GC) gas chromatography equipped with a TCD detector and a Carboxen column. Methane and other hydrocarbons were detected by a flame ionization detector (FID) with Alumina S column.

Subsequently, the catalysts whose performance was measured in the GC test analysis system were tested on three different parameters. Experiments were carried out by changing the pressure, temperature and flow parameters from the reaction conditions. Conditions of 200 °C, 250 °C, 300 °C and 350 °C for temperature, 0.08, 0.16, 0.32 and 0.64 for flow and 1, 3, 5 and 10 bar for pressure were applied (Table 1).

Table 1. Reaction conditions.

Pressure (bar)	Temperature (°C)	Flow (L/min)
1	200	0.08
3	250	0.16
5	300	0.32
10	350	0.64

3. Characterization Analysis

The elementary compositions of samples were determined using a; Thermo Scientific, ARL PERFORM'X Sequential PFX-374 analyzer used as X-ray source.

A quantitative analysis of the coke amount on the used catalyst was performed with a thermogravimetry analyzer (TGA, Perkin Elmer Pyris1). Thermogravimetric analyzes are performed in a one-step temperature program. Heating is 900 °C with a heating rate of 10°C/min under 30 mL/min air flow. Approximately 10 mg of sample is used for TGA analysis.

The physical structure of the surface of catalyst was studied by scanning electron microscopy (SEM, JEOL-6390LV). The samples were coated with a Pd-Au mixture prior to analysis. SEM analysis method is an important technique for surface topography studies. In this analysis method, it provides information about the shape, size, position and morphology of metal particles on the support surface.

There are many commercially available catalysts that are utilized to produce methane from syngas. The content of one of them, was determined by XRF characterization analysis (Table 2). As can be seen from the XRF analysis results in Table 2, most of the commercial catalyst structure consists of nickel oxide. Apart from this, Al_2O_3 , SiO_2 , MgO , CuO , La_2O_3 etc. metal salts are also present in the structure.

Table 2. XRF analysis results.

Compounds	% Content	Compounds	% Content
NiO	70.58	MoO ₃	0.0068
Al ₂ O ₃	11.47	Tb ₄ O ₇	0.0067
SiO ₂	08.69	K ₂ O	0.0064
MgO	5.4	ZrO ₂	0.0053
CuO	1.5	IrO ₂	0.0053
La ₂ O ₃	1.23	Sc ₂ O ₃	0.0048
Ho ₂ O ₃	0.571	Ga ₂ O ₃	0.004
Fe ₂ O ₃	0.104	Cr ₂ O ₃	0.004
CaO	0.0923	SnO ₂	0.0037
Na ₂ O	0.0669	Au	0.0033
ZnO	0.0484	CeO ₂	0.0022
BaO	0.0452	RuO ₄	0.0021
Dy ₂ O ₃	0.0411	Tl ₂ O ₃	0.0017
S	0.0286	SeO ₂	0.0013
TiO ₂	0.0207	Br	0.00084
P ₂ O ₅	0.0188	SrO	0.0005
Y ₂ O ₃	0.0178	PbO	0.0004
Ag ₂ O	0.0134	ThO ₂	0.0004
Re ₂ O ₇	0.0083	GeO ₂	0.0003

To explore the presence of carbon deposits on the catalyst's surface, TGA and SEM analyses were performed on spent catalysts.

The extent of coke formation on the catalyst after the reaction was quantitatively studied by TGA in an oxidative atmosphere, and the results are shown in Figure 1. Figure 1 shows that there was no significant weight loss due to the C-formation. From the TGA analysis of the used catalyst, no clear evidence of carbon on the catalyst was found. This result indicates that the catalyst has the high resistance to the formation of coke.

The fresh and used catalysts were characterized with scanning electron microscopy. Figure 2 (a–f) illustrates the results of scanning electron microscope of the same samples. The coke formation that is expected to occur in catalysts containing Ni particles is the formation of fibrous carbon. Between the Ni particles and the support, fibrous carbon forms and grows, which leads to clogging of the catalyst pores [14]. Figure 2 (d,e,f) SEM images show that the surface of the spent catalysts is almost smooth. The surface properties of used catalysts are very similar to those of fresh catalysts. Any fibrous structure is not observed on the surface.

Another distinctive characteristic observed in the SEM micrographs is the prevalence of bright areas is indicator of exposed nickel metal. Figure 2, the SEM images of fresh (b) and spent (e) catalysts at the same magnification ratios show similar intensity of bright areas. This indicates that the exposed active nickel surfaces have similar amounts and are not affected by coke formation [15].

The SEM and the TGA results confirmed each other and according to these results zero coke formation was recorded for this catalyst.

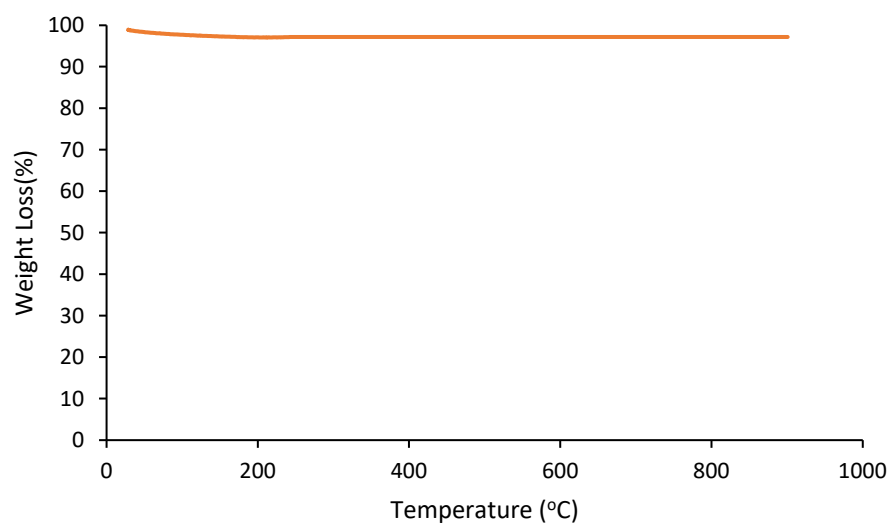


Figure 1. Thermogravimetric profile of used catalyst.

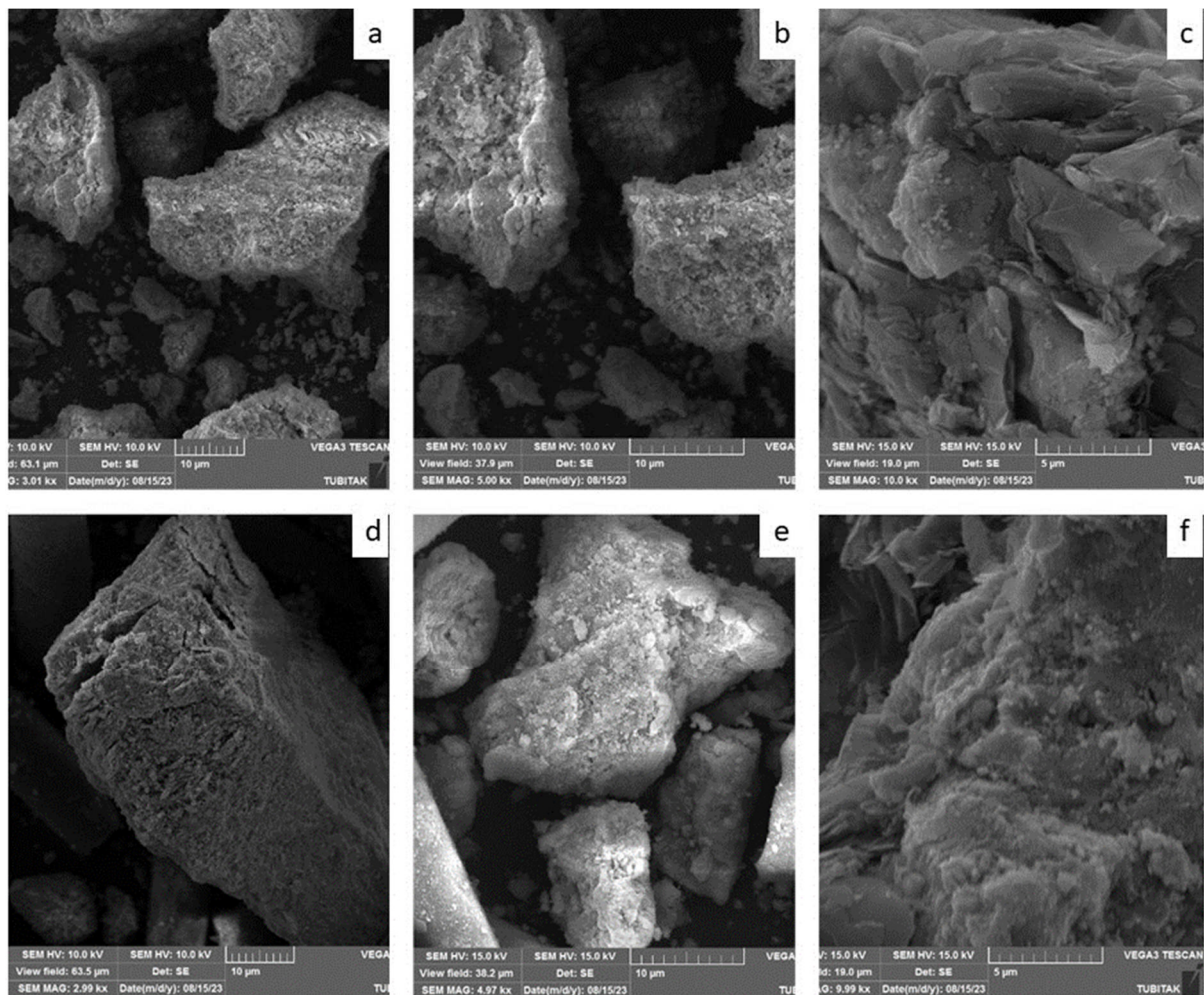


Figure 2. SEM of the fresh catalyst (a) 3kx, (b) 5kx, (c)10kx. SEM of the spent catalyst (d) 3kx, (e) 5kx, (f)10kx. .

The hydrogen consumption versus time graph was generated with the data obtained (Figure 3). It was observed that the reduction was complete.

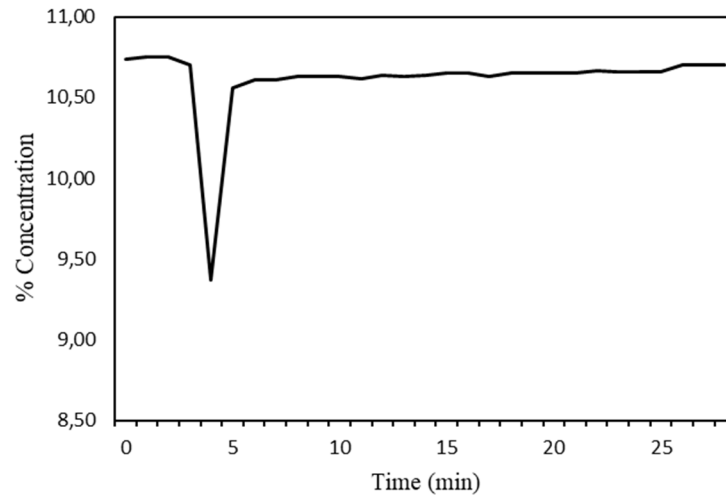


Figure 3. Hydrogen consumption in CO₂ methanation.

4. Methods

4.1. Decision Tree Regression (DTR)

A decision tree is a mostly employed technique in machine learning for creating predictive models based on simple logical statements [16]. The fundamental decision tree structure involves a starting point (root node), intermediate points (internal nodes), and end points (terminal nodes) [17,18]. In the context of predicting continuous variables rather than categorical ones, Decision Tree Regression (DTR) is utilized. DTR progressively divides the dataset into smaller groups to achieve better uniformity within these groups. Subsequently, it aims to minimize predictive errors within each of these divisions. This process ultimately generates predictions for continuous output variables through logical rules. In decision tree algorithms, the arrangement of these groups is determined by evaluating the information gain for all feasible groupings and selecting the grouping with the highest information gain to serve as the root node, as described in Equation (2) [19].

$$Gain(C, A) = S(C) - \sum_{v \in A} \frac{|C_v|}{|C|} S(C_v) \quad (1)$$

In the given equation, where C represents the cluster set, A denotes the attribute, and $S(C)$ stands for the entropy of cluster set C . The primary strength of Decision Tree Regression (DTR) over alternative modeling methods lies in its capability to generate a model that is expressed as a collection of logical rules and expressions. Moreover, it can effectively handle scenarios involving multiple output targets. Furthermore, it possesses the advantage of being easily comprehensible and interpretable, allowing for visualization of the tree structures [20]. Notwithstanding these noteworthy qualities, decision trees come with certain drawbacks. They are not designed to accommodate outputs reliant on more than one feature, they can produce somewhat inconsistent outcomes, they exhibit high sensitivity to minor alterations in test data, and they tend to generate intricate tree configurations when dealing with numerical datasets [17].

The evolution of utilizing decision trees in classification and estimation involves the creation of a decision tree model based on training data, the assessment of this model using test data and appropriate evaluation criteria, and the application of the resultant model to predict future values [18].

4.2. Random Forest Regression

The method of random forest regression (RF) entails the fitting of numerous classification trees to a given dataset. Subsequently, the predictions generated by these individual trees are amalgamated to construct a final predictive model, as explained by Golden et al. in 2019 [21]. Unlike a solitary model, the random forest approach leverages the combined predictions of multiple models

to yield more robust and broadly applicable outcomes. Various ensemble techniques, including bagging and boosting, involve the utilization of multiple classifiers in a single classification task. These approaches amalgamate the outputs of these classifiers through methods such as voting or averaging to enhance predictive efficacy. Bagging, or bootstrap aggregation, is one ensemble method that entails training multiple models independently and concurrently. The resulting predictions are then combined to establish an averaged model, as elaborated by Soni et al. in 2021 [22]. Boosting, on the other hand, is a distinct ensemble method involving the sequential and adaptive training of weak learners to enhance the predictions of a learning algorithm. Within the context of the RF method discussed in this study, the least-squares boosting algorithm was employed. This algorithm involves fitting the residual errors of individual weak learners to achieve superior performance.

4.3. Gaussian Process Regression

GPR represents a machine learning technique. This specific algorithm is kernel-centered, not relying on fixed parameters, and has a foundation in probability [23]. Regression models utilizing Gaussian processes are entirely rooted in probability, adaptable, and straightforward to employ, leading to their widespread use across various applications. However, its drawback lies in its effectiveness being constrained to a relatively small number of cases, usually no more than a few thousand. This limitation stems from the rapid increase in computational and memory demands, growing cubically and quadratically with the number of training cases, respectively [24].

A Gaussian process is fully defined by its mean function, denoted as $m(x)$, and covariance function, represented as $k(x, x')$. These functions describe a real process $f(x)$, as shown in Equation (3) [25,26]. When appropriately scaled, the mean function $m(x)$ can often be set to zero [25,27].

$$f(x) \sim GP(m(x), k(x, x')) \quad (3)$$

An algorithm for application of GPR is as follows [25]:

1. **Input:** X (inputs), y (targets), σ^2 (noise level), x^* (test input).
2. $L := \text{cholesky}(K + \sigma^2 I)$
3. $\alpha := L^T \backslash (L \backslash y)$
4. $\bar{f}^* := k_*^T \alpha$
5. $v := L \backslash k^*$
6. $\mathbb{V}[f^*] := k(x^*, x^*) - v^T v$
7. $\log(p(y|X)) := -\frac{1}{2} y^T \alpha - \sum_i \log(L_{ii}) - \frac{n}{2} \log(2\pi)$
8. **Return:** \bar{f}^* (mean), $\mathbb{V}[f^*]$ (variance), $\log(p(y|X))$ (log marginal likelihood)

4.4. Support Vector Machines

Support Vector Machines (SVM) have gained popularity in the realm of machine learning due to their non-parametric structure and incorporation of kernel functions. Among SVM techniques, SVM stands out as a method tailored for predicting continuous variables. Differing from polynomial regression and Artificial Neural Networks (ANNs), SVM models are trained by minimizing the bound on generalization error, as opposed to prediction error [28]. Thus, SVM models prioritize achieving optimal generalization performance.

The fundamental principle of SVM involves mapping data to a higher-dimensional feature space using a kernel function. Within this space, an optimal hyperplane is sought, aiming to minimize a chosen cost function [29]. Consequently, SVM conducts linear regression within an expanded dimensionality using a kernel function (denoted as K) as depicted in Equation (4).

In contrast, ANNs possess a distinctive capability to predict multiple output variables using a single model. In the context of this study, where the dataset contains four features and four corresponding output variables, a singular ANN model was devised. This model comprises two hidden layers, each containing nine neurons, as illustrated in Figure 1. The hidden layers employ sigmoid activation functions, while the output layer employs a linear activation function. The training of the neural network involves adjusting connection weights to minimize the disparity

between predictions and actual outputs, with the widely used backpropagation algorithm employed for this purpose [30].

Top of Form

$$\hat{y} = \sum \beta_i * K(n \ i = 1 \ x_i, x) + c \quad (4)$$

In the given context, the symbols β_i and c represent constant coefficients, while $(, x)$ stands for the kernel function. Within this study, a kernel function in the form of a second-degree polynomial (as indicated by Equation 5) was employed. This specific kernel was chosen due to the limitations observed with linear and Gaussian kernels in capturing the intricacies of the problem at hand. Both linear and Gaussian kernels exhibited insufficient predictive capabilities, prompting the exploration of alternative kernels.

While the use of third-degree or higher polynomial kernels could potentially enhance performance, their adoption was deemed impractical due to substantial computational overhead. The marginal performance gain achieved by these higher-degree kernels did not justify the extensive training time they demanded.

Throughout the training phase of the models, the epsilon-insensitive cost function (expressed through Equations 6 and 7) was utilized and subsequently minimized.

Top of Form

$$(x_i, x) = (x_i'x + 1)^2 \quad (5)$$

$$L_\epsilon = \begin{cases} 0 & \text{if } |y - \hat{y}| \leq \epsilon \\ |y - \hat{y}| - \epsilon & \text{otherwise} \end{cases} \quad (6)$$

$$J = \sum_{i=1}^n L_\epsilon(\hat{y}_i, y) + \frac{\lambda}{2} |w|^2 \quad (7)$$

4.5. Validation Method: K-fold Cross Validation (KCV)

Cross-validation methods are used to objectively evaluate the performance of a prediction model. K-fold cross-validation (KCV) is a highly preferred cross-validation technique due to its use of the entire data set [31]. In KCV, the data is divided into k equally sized partitions (folds). The candidate model is then trained on training folds and evaluated on the remaining test fold [32]. This process is repeated k times with a different fold as the test set each time, ensuring that all data is used for both training and testing. In this study, 'K' is chosen as ten because it has been shown to reflect the most unbiased generalization performance of the models compared to other choices. This approach reduces variance by using 10 different partitions, and the performance estimate is less sensitive to the random partitioning of the data compared to other cross-validation methods [33].

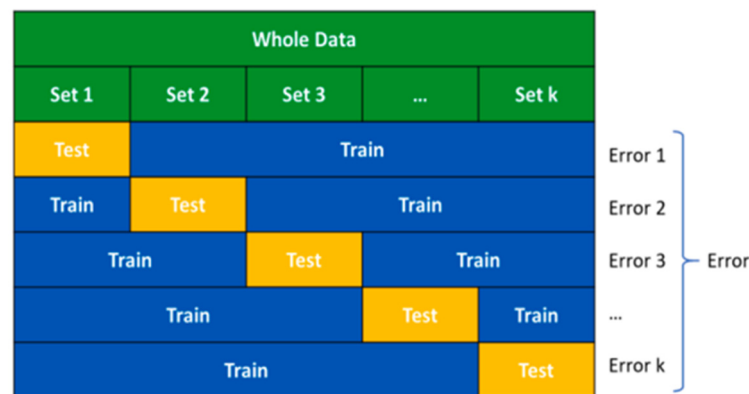


Figure 4. The principles of k-fold cross validation.

5. Results & Discussion

All prediction models are trained to predict CO₂ methanation output concentrations and their performances are evaluated according to their test performance in 10-fold cross validation technique. R-squared (R²) is a statistical measure that represents the proportion of variance in the dependent variable (outcome) that's explained by the independent variables (predictors) in a regression model. It is also called the coefficient of determination. R-squared is used to assess the goodness of fit of a regression model and how well the model's predictions match the actual observed data. Adjusted R² shows how well the terms fit a line, but adjusts for the number of terms in a pattern. Root mean square error (RMSE) measures the average difference between a statistical model's predicted values and the actual values. Normalized Root Mean Square Error (NRMSE) calculates relative error between predicted and observed values [34]. Absolute Average Error (AAE) is a measure of the average size of the mistakes in a collection of predictions, without taking their direction into account.

$$R^2 = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y})^2} \quad (8)$$

$$Adj. R^2 = 1 - \left[\frac{(1-R^2) * (m-1)}{m-n-1} \right] \quad (9)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{m}} \quad (10)$$

$$NRMSE = \frac{\sqrt{\frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{m}}}{\bar{y}_i} \quad (11)$$

$$AAE = \frac{1}{n} \sum_{i=1}^n |x_i - \hat{x}| \quad (12)$$

When analyzing the data in Table 3, the R² value indicating linearity is expected to be closest to 1 and the RMSE value indicating the error rate is expected to be the smallest. Looking at Table 3, it is seen that the RF method is the closest to 1 with an R² value of 0.987 for CO₂ data. This is supported by the smallest RMSE value of 0.526. The RF method was determined to be the most appropriate method for CO₂ conversion data. When the H₂ conversion data is analyzed, it is seen that the RF method has the closest R² value to 1 with 0.987. Although the error rate of 2.292 is very large, it is the smallest among the RMSE values and the most appropriate method is determined as RF method. In N₂ data, the closest R² value to 1 is seen in the DT method with 0.986. The smallest RME value is 1.392, supporting the selection of DT as the most appropriate method. In the CH₄ data, the largest R² value is seen in the DT method and this is supported by the RMSE value of 1.640. As a result, RF method selection in CO₂ and H₂ data, DT method selection was determined as the most suitable for processing N₂ and CH₄ data.

Table 3. Machine learning results.

		R²	aR²	RMSE	NRMSE	AAE
CO₂	RF	0.987	0.986	0.526	0.068	0.033
	DT	0.985	0.985	0.552	0.073	0.039
	SVM	0.825	0.821	1.900	0.242	0.176
	GP	0.950	0.948	1.020	0.142	0.096
H₂	RF	0.987	0.987	2.292	0.110	0.056
	DT	0.985	0.984	2.539	0.093	0.057
	SVM	0.818	0.813	8.708	0.451	0.287
	GP	0.929	0.927	5.430	0.314	0.191
N₂	RF	0.975	0.975	1.836	0.047	0.028
	DT	0.986	0.986	1.392	0.039	0.026
	SVM	0.789	0.784	5.384	0.172	0.136

CH ₄	GP	0.918	0.916	3.351	0.086	0.069
	RF	0.980	0.980	1.750	0.909	0.355
	DT	0.983	0.982	1.640	1.448	0.675
	SVM	0.807	0.802	5.485	33.617	10.743
	GP	0.929	0.927	3.322	3.866	1.754

When Figure 5 is analyzed, it is seen that the data in SVM and GO methods are too much out of the line. The data outside the line are more than the data in RF and DT methods. It is also seen that there is clustering in the data in DT, SVM and GP methods. This proves that these models are more inappropriate than RF. For these reasons, it is decided that RF is the most appropriate model.

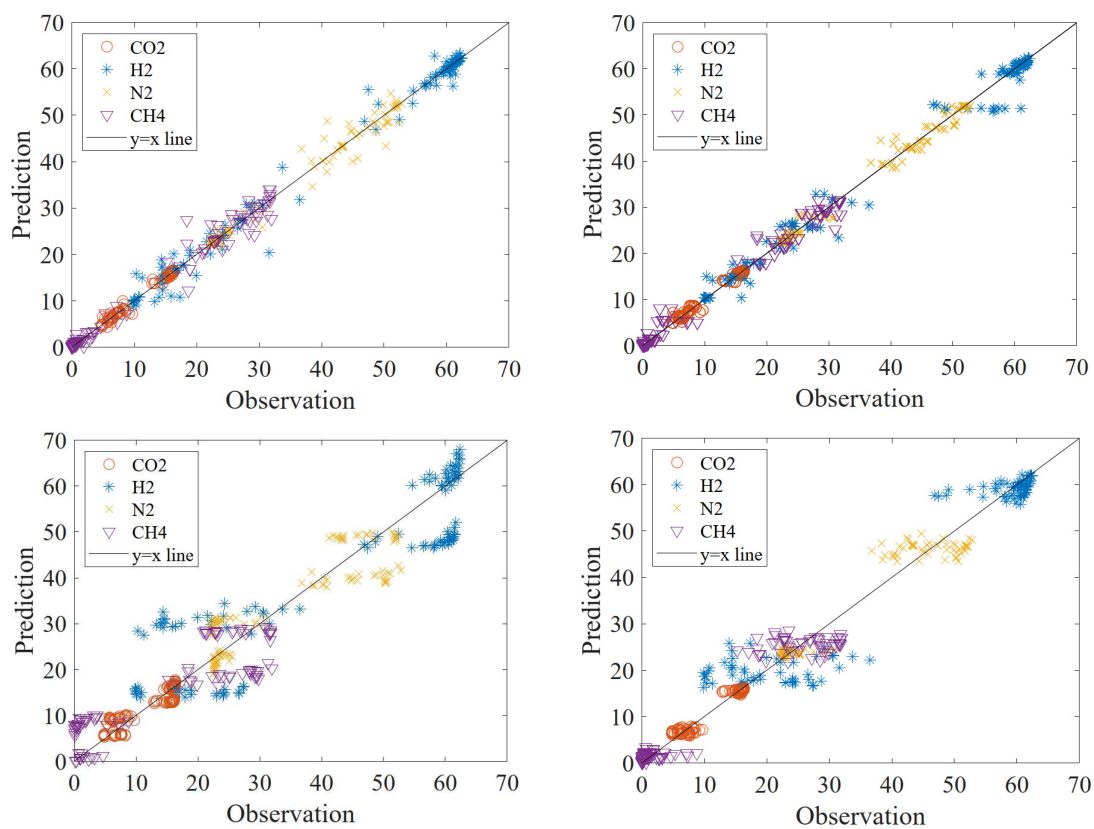


Figure 5. The observed vs. predicted gas concentrations (a)RF (b)DT (c)SVM (d)GP.

When Table 4 is examined, R² values close to 1 are important for the accuracy of the data. In the 4th research article, it is seen that the train R² value is smaller than the test R² value. This situation is far from reality. Although there is a lot of data in research articles 1 and 7, it is seen that the R² values of the tests are far from 1. In research paper 10, the number of data is close to the number of data in this study, but it is seen that it does not reach the R² value we have reached. Among the R² values in the literature studies, this study has the closest R² value to 1. It is seen that the best result among the literature studies is reached in this study.

Table 4. Comparison with the best performing algorithms in the literature.

Data Size	Prediction model	Validation method (Test/ Train)	RMSE	R ²	Ref. no
4051	Random forest (RF)	80/20	6.4 12.7	0.97 (train) 0.85 (test)	[1]
136	Genetic programming (GP)	85/15	-	0.94 (train) 0.95 (test)	[4]

1425	Extreme gradient boost (XGB)	85/15	0.09	0.98 (train) 0.88 (test)	[7]
4360	XGBoost	10-fold cross-validation	7.27	0.95	[8]
93	Extreme learning machine (ELM)	5 Times Repeated Hold out 90/10	0.18	0.87	[10]
116	Random Forest	10-fold cross-validation	0.53	0.99	This Study

6. Conclusions

SEM and TGA techniques were used to both qualitatively and quantitatively analyze the presence of carbon deposition. From the TGA graph the spent catalyst displayed no weight loss in the spent catalyst, indicating an absence of coke. TGA and SEM results confirmed each other and according to these results zero coke formation was recorded for this catalyst.

The RF method was the most successful method with an R^2 value of 0.987 for CO_2 data. The method was also successful to predict other gas concentrations. The results were compared with the literature. It is seen that the best result among the literature studies is reached in this study. With this study, we show that the use of machine learning methods is capable of predicting output concentration of CO_2 methanation process. The RF methods have a low computational requirement to make predictions once it is trained. Therefore, this study offers more flexibility in controlling and optimizing conversion of CO_2 .

References

1. Yilmaz, B., Oral, B., & Yıldırım, R. (2023). Machine learning analysis of catalytic CO_2 methanation. *International Journal of Hydrogen Energy*, 48(64), 24904-24914. <https://doi.org/10.1016/j.ijhydene.2022.12.197>
2. Hong, T., Saifuddin, N., Halim, S., Young-Kwon, P., Hernández, & H., Loke, S. (2022). "Current Developments in Catalytic Methanation of Carbon Dioxide—A Review" *Frontiers in Energy Research* Vol. 9 <https://doi.org/10.3389/fenrg.2021.795423>.
3. Kim, C., Yoo, C., Oh, H., Min, B., & Lee, U. (2022). "Review of carbon dioxide utilization technologies and their potential for industrial application", *Journal of CO_2 Utilization*, Volume 65, <https://doi.org/10.1016/j.jcou.2022.102239>.
4. Dashti, A., Bazdar, M., Fakhrabadi, E., & Mohammadi, A. (2022). Modeling of Catalytic CO_2 Methanation Using Smart Computational Schemes. *Chemical Engineering & Technology*, 45, 135-143. <https://doi.org/10.1002/ceat.202100557>
5. Fedorov, A., & Linke, D. (2022). Data analysis of CO_2 hydrogenation catalysts for hydrocarbon production. *Journal of CO_2 Utilization*, 61. <https://doi.org/10.1016/j.jcou.2022.102034>
6. Ngo, S. I., & Lim, Y.-I. (2021). Solution and Parameter Identification of a Fixed-Bed Reactor Model for Catalytic CO_2 Methanation Using Physics-Informed Neural Networks. *Catalysts*, 11(11), 1304. <http://dx.doi.org/10.3390/catal11111304>
7. Suvarna, M., Araújo, T. P., & Pérez-Ramírez, J. (2022). A generalized machine learning framework to predict the space-time yield of methanol from thermocatalytic CO_2 hydrogenation. *Applied Catalysis B: Environmental*, 315. <https://doi.org/10.1016/j.apcatb.2022.121530>
8. Chattoraj, J., Hamadicharef, B., Kong, J., Pargi, M. K., Zeng, Y., Poh, C. K., Chen, L., Gao, F., & Tan, T. L. (2022). Theory-Guided Machine Learning to Predict the Performance of Noble Metal Catalysts in the Water-Gas Shift Reaction. *ChemCatChem*, 14.
9. Pérez, S., Aragón, J. J., Pecina, I., et al. (2019). Enhanced CO_2 Methanation by New Microstructured Reactor Concept and Design. *Topics in Catalysis*, 62, 518–523. <https://doi.org/10.1007/s11244-019-01139-4>
10. Shen, Y., Dong, Y., Han, X., Wu, J., Xue, K., Jin, M., Xie, G., & Xu, X. (2023). Prediction model for methanation reaction conditions based on a state transition simulated annealing algorithm optimized extreme learning machine. *International Journal of Hydrogen Energy*, 48, 24560-24573. <https://doi.org/10.1016/j.ijhydene.2022.10.031>
11. Praveen, C. S., & Comas-Vives, A. (2020). Design of an Accurate Machine Learning Algorithm to Predict the Binding Energies of Several Adsorbates on Multiple Sites of Metal Surfaces. *ChemCatChem*, 12.
12. Yeo, C.-E., Seo, M., Kim, D., Jeong, C., Shin, H.-S., & Kim, S. (2021). Optimization of Operating Conditions for CO_2 Methanation Process Using Design of Experiments. *Energies*, 14(24), 8414. <http://dx.doi.org/10.3390/en14248414>

13. Kristiani, A., & Takeishi, K. (2022). CO₂ methanation over nickel-based catalyst supported on yttria-stabilized zirconia. *Catalysis Communications*, 165. <https://doi.org/10.1016/j.catcom.2022.106435>
14. Song, D. H., Jung, U. H., Kim, Y. E., Im, H. B., Lee, T. H., Lee, K. B., & Koo, K. Y. (2022). Influence of Supports on the Catalytic Activity and Coke Resistance of Ni Catalyst in Dry Reforming of Methane. *Catalysts*, 12(2). <https://doi.org/10.3390/catal12020216>
15. Venugopal, A., Naveen Kumar, S., Ashok, J., Hari Prasad, D., Durga Kumari, V., Prasad, K. B. S., & Subrahmanyam, M. (2007). Hydrogen production by catalytic decomposition of methane over Ni / SiO₂. *International Journal of Hydrogen Energy*, 32(12), 1782–1788. <https://doi.org/10.1016/j.ijhydene.2007.01.007>
16. Özkan, K., Işık, Ş., Günkaya, Z., Özkan, A., & Banar, M. (2019). A heating value estimation of refuse derived fuel using the genetic programming model. *Waste Management*, 100, 327-335.
17. Rokach, L., & Maimon, O. (2010). Classification Trees. In O. Maimon & L. Rokach (Eds.), *Data Mining and Knowledge Discovery Handbook* (pp. 149-174). Springer US. https://doi.org/10.1007/978-0-387-09823-4_9
18. Zhao, Y., & Zhang, Y. (2008). Comparison of decision tree methods for finding active objects. *Advances in Space Research*, 41, 1955-1959. <https://doi.org/10.1016/j.asr.2007.07.020>
19. Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1, 81-106. <https://doi.org/10.1007/BF00116251>
20. Elmaz, F., Yücel, Ö., & Mutlu, A. Y. (2020). Machine learning based approach for predicting of higher heating values of solid fuels using proximity and ultimate analysis. *International Journal of Advances in Engineering and Pure Sciences*, 32(2), 145-151.
21. Golden, C. E., Rothrock, M. J., Jr., & Mishra, A. (2019). Comparison between random forest and gradient boosting machine methods for predicting *Listeria* spp. prevalence in the environment of pastured poultry farms. *Food Research International*, 122, 47-55. <https://doi.org/10.1016/j.foodres.2019.03.062>
22. Soni, A., Al-Sarayreh, M., Reis, M. M., & Brightwell, G. (2021). Hyperspectral imaging and deep learning for quantification of *Clostridium sporogenes* spores in food products using 1D-convolutional neural networks and random forest model. *Food Research International*.
23. Makridakis, S., Spiliotis, E., Assimakopoulos, V., & Hernandez Montoya, A. R. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLoS One*, 13(3), e0194889.
24. Candela, J. Q., & Rasmussen, C. (2005). A Unifying View of Sparse Approximate Gaussian Process Regression. *Journal of Machine Learning Research*, 6, 1939-1959.
25. Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. Massachusetts Institute of Technology.
26. Mathworks. (2021). Gaussian Process Regression Models. <https://www.mathworks.com/help/stats/gaussian-process-regression-models.html>
27. Birgen, C., Magnanelli, E., Carlsson, P., Skreiberg, Ø., Mosby, J., & Becidan, M. (2021). Machine learning based modelling for lower heating value prediction of municipal solid waste. *Fuel*, 283, 118906. <https://doi.org/10.1016/j.fuel.2020.118906>
28. Evgeniou, T., Pontil, M., & Poggio, T. (2000). Regularization Networks and Support Vector Machines. *Advances in Computational Mathematics*, 13(1), 1-50.
29. Hecht-Nielsen, R. (1988). Theory of the backpropagation neural network. *Neural Networks*, 1(1), 445.
30. Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, 14(3), 199-222.
31. Elmaz, F., Yücel, Ö., & Mutlu, A. Y. (2019). Evaluating the effect of blending ratio on the co-gasification of high ash coal and biomass in a fluidized bed gasifier using machine learning. *Mugla Journal of Science and Technology*, 5(1), 1-12.
32. Büyükçakır B, Mutlu AY. Comparison of hilbert vibration decomposition with empirical mode decomposition for classifying epileptic seizures. In: 2018 52nd asilomar conference on signals, systems, and computers. IEEE; 2018. p. 357e62.
33. Ambroise C, McLachlan GJ. Selection bias in gene extraction on the basis of microarray gene-expression data. *Proc Natl Acad Sci* 2002;99(10):6562e6.
34. Yaka, H., Insel, M. A., Yucel, O., & Sadikoglu, H. (2022). A comparison of machine learning algorithms for estimation of higher heating values of biomass and fossil fuels from ultimate analysis. *Fuel*, 320, 123971.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.